Plant Physiol. (1992) 100, 582–592
0032-0889/92/100/0582/11/$01.00/0

Received for publication June 6, 1992
Accepted June 16, 1992

A Phosphoribosylanthranilate Transferase Gene Is Defective in Blue Fluorescent *Arabidopsis thaliana* Tryptophan Mutants[1]

Alan B. Rose, Amy L. Casselman, and Robert L. Last*

Boyce Thompson Institute for Plant Research and Section of Genetics and Development, Cornell University, Tower Road, Ithaca, New York 14853–1801

ABSTRACT

An *Arabidopsis thaliana* gene encoding phosphoribosylanthranilate transferase is shown to be the gene that is defective in blue fluorescent *trp1* mutant plants. This gene, named *PAT1*, was isolated using an *A. thaliana* cDNA clone that suppressed an *Escherichia coli trpD*⁻ mutation. The *PAT1* coding region is homologous to those for the phosphoribosylanthranilate transferases from many microorganisms. Unlike other genes involved in aromatic amino acid biosynthesis in *A. thaliana*, *PAT1* appears to be a single-copy gene. *PAT1* was demonstrated to be the gene that is defective in blue fluorescent *trp1* mutants by two methods: genetic complementation in transgenic plants and genetic mapping studies. This is the first report of cloning a plant phosphoribosylanthranilate transferase gene. The *PAT1* gene should prove useful as a selectable marker for transformation or a visible reporter of gene expression when used in conjunction with *trp1* plants.

The tryptophan biosynthetic pathway is the source of many important compounds in plants. In addition to its role in protein synthesis, this pathway provides precursors for a variety of indolic secondary products. These include the plant growth regulator auxin (IAA) (36), antimicrobial phytoalexins (32), alkaloids, and glucosinolates (15).

Relatively little is known about the regulation of tryptophan biosynthesis in plants, although this pathway has been studied extensively in microorganisms. The pathway itself, shown in Figure 1, is identical in all microorganisms studied to date, and the hypothesis that plants will utilize the same set of reactions is supported by biochemical analysis (reviewed in refs. 3 and 30). Also, all of the existing plant tryptophan mutants have biochemical lesions in steps found in the microbial pathway (21, 22, 36). However, a definitive statement that Figure 1 represents the in vivo plant pathway will not be possible until all steps are defined by mutation and/or analysis of the plant genes.

The tryptophan-requiring *trp1-1* mutant of *Arabidopsis thaliana* was the first described plant amino acid auxotroph in which the auxotrophy shows Mendelian inheritance (22). Biochemical analysis revealed that *trp1* plants are defective in PAT[2] (EC 2.4.2.18) activity, and are therefore unable to convert anthranilate to PR-anthranilate. These plants have a number of interesting and potentially useful characteristics. For example, the *trp1* plants are blue fluorescent under UV light because they accumulate anthranilate compounds. The *trp1-1* plants are also small and bushy, consistent with a defect in IAA biosynthesis.

We have isolated and characterized the *A. thaliana PAT1* gene, the first PR-anthranilate transferase gene from a plant to be cloned and analyzed. *PAT1* appears to be a single-copy gene, unlike the previously described aromatic amino acid biosynthetic genes. The *PAT1* gene complements and is genetically linked to *trp1* mutations, consistent with the hypothesis that *trp1* alleles are *PAT1* structural gene defects. The *PAT1* gene, together with the *trp1* plants, will add to the growing collection of molecular tools available for the manipulation and study of *A. thaliana*.

MATERIALS AND METHODS

Plant Methods

Arabidopsis thaliana ecotype Columbia was grown on agar medium or soilless mix as described (21, 22). Standard methods were used to isolate plant DNA (6) and total RNA (12).

DNA Manipulations

Plasmids were constructed using enzymes from New England Biolabs or Promega according to the recommendations of the manufacturers. *Escherichia coli* strains JM109 and NM538 were used for plasmid and bacteriophage lambda propagation, respectively. Single-stranded DNA was sequenced using the Multiwell Microtitre Plate DNA Sequencing System (Amersham).

pNN423, which contains an *A. thaliana* cDNA capable of suppressing an *E. coli trpD*⁻ mutation, was isolated as previously described by Elledge and coworkers (13) and was provided by John Mulligan (Stanford University). The entire cDNA insert in pNN423 was sequenced by subcloning re-

[1] Supported by National Institute of Health grant GM43134 and National Science Foundation Presidential Young Investigator Award DMB-9058134 to R.L.L.

[2] Abbreviations: PAT, phosphoribosylanthranilate transferase (although this enzyme is frequently called antranilate phosphoribosyl transferase, we have chosen the designation PAT to avoid confusion with the *A. thaliana APT1* gene, which encodes adenine phosphoribosyl transferase [24]); PR-anthranilate, phosphoribosylanthranilate; PNS, plant nutrient medium with 0.5% (w/v) sucrose; RFLP, restriction fragment-length polymorphism.

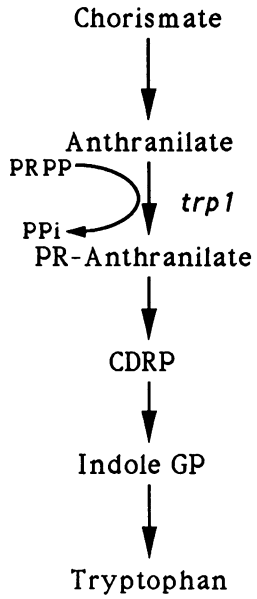

Figure 1. The tryptophan biosynthetic pathway. The enzyme defective in *trp1* mutant plants (PAT) is indicated. Abbreviations: PRPP, 5-phosphoribosyl-1-pyrophosphate; PPi, inorganic pyrophosphate; PR-anthranilate, 5-phosphoribosylanthranilate; CDRP, 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate; Indole GP, indole-3-glycerol phosphate.

striction fragments into pBluescript KS(+) (Stratagene). We also determined the sequence of several hundred nucleotides from the ends of three other *PAT1* cDNAs isolated as plasmids that suppress *E. coli trpD*$^-$. The GenBank and EMBL protein data bases were searched using the FASTA algorithm (27).

A genomic clone (λAR1–5) that was obtained from an *A. thaliana* library (Clonetech) in λEMBL3 by plaque hybridization using pNN423 as a probe (28) was characterized. Fragments that cross-hybridize to the *PAT1* cDNA were subcloned as follows. pAR119, which contains an 8.6-kb *Sal*I fragment cloned into KS(+), was digested with *Bam*HI, releasing a 4.1-kb *Bam*HI-*Sal*I genomic fragment with 45 bp of the KS(+) polylinker. pAR125 and pAR134 contain this 4.1-kb *Bam*HI fragment in pGEM7Zf(+) (Promega) in opposite orientations. For sequence analysis, nested unidirectional deletions were generated from *Cla*I- and *Kpn*I-digested pAR125 and pAR134 DNA by a method modified from that of Henikoff (17), which involves treatment with exonuclease III followed by mung bean nuclease (New England Biolabs). The complete sequence of both strands of the 4.1-kb fragment was determined and submitted to GenBank (accession No. M96073).

Plasmid pAR134Δ74, a deletion derivative of pAR134 that retains *PAT1* nucleotides 1 to 1247 (numbering as in Fig. 5), was used to generate the probe for RNase protection experiments. The primer extension control plasmid pAR163 contains a 253-bp *Sca*I-*Xho*I fragment from pAR134Δ74, includ-

Figure 2. Phenotypic properties of *trp1–100* plants. A, *trp1–100* plants are prototrophic. Wild-type (left), *trp1–100* (center), and *trp1–1* (right) plants were grown on PNS medium lacking tryptophan for 15 d at a light level of 150 μmol PAR·m^{-2}·s^{-1}. B, Adult morphology of *trp1–100* plants. From left to right, 4-week-old wild type, *trp1–100*, *trp1–1*, and 9-week-old *trp1–1* plants grown in soil are shown.

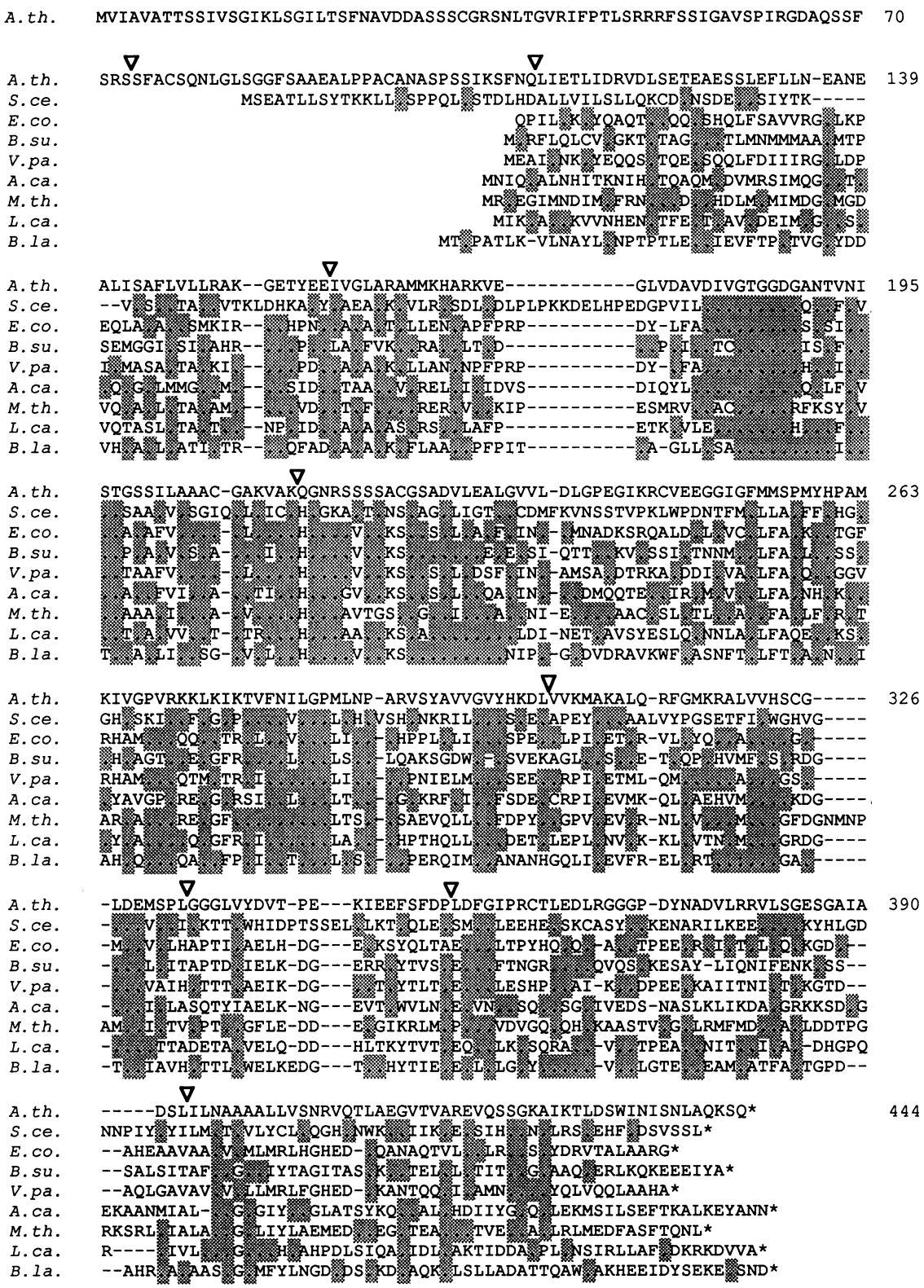

Figure 3. Comparison of the PAT1 protein to other PATs. The protein (single-letter code) predicted to be encoded by the *PAT1* cDNA from *A. thaliana* (*A.th.*) was aligned using the PILEUP program of the University of Wisconsin GCG package to the PATs from *Saccharomyces cerevisiae* (*S.ce.*), *E. coli* (*E.co.*), *Bacillus subtilis* (*B.su.*), *Vibrio parahaemolyticus* (*V.pa.*), *Acinetobacter calcoaceticus* (*A.ca.*), *Methanobacterium thermoautotrophicum* (*M.th.*), *Lactobacillus casei* (*L.ca.*), and *Brevibacterium lactofermentum* (*B.la.*). Residues identical to the *A. thaliana* sequence are indicated by shaded dots. Gaps introduced to optimize the alignment are represented by dashes, asterisks indicate stop codons, and inverted triangles show the positions of introns in the *A. thaliana* gene.

ing *PAT1* nucleotides 1005 to 1247 and 10 bp of linker, in the *Eco*RV-*Xho*I sites of KS(+).

DNA hybridization probes were generated using the Megaprime DNA Labeling Kit (Amersham) with [α-^{32}P]dCTP (New England Nuclear-Dupont), and unincorporated nucleotides were removed with Nuctrap push columns (Stratagene). DNA was transferred to Zeta-Probe filters (Bio-Rad) by alkaline transfer according to the manufacturer's directions. Standard stringency hybridizations and washes were carried out as described (5).

RNA Analysis

The oligonucleotide ARP1 (5′-GGAGCTCGTCGTCG-CCACCGC-3′), which is complementary to nucleotides 1176 to 1196, was 5′ end labeled by polynucleotide kinase and used for primer extension analysis as described (1). Approximately 5×10^5 cpm of primer were annealed to 10 μg of RNA overnight at 48°C and extended at 42°C using 100 units of Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories). Control sense RNA was synthesized in vitro from *Xho*I-digested pAR163 DNA using T7 RNA polymerase (Promega).

RNase protection experiments were performed as detailed (1). Riboprobes were synthesized in vitro from *Sca*I-digested pAR134Δ74 using T7 RNA polymerase. Approximately 5×10^7 cpm of gel-purified probe were hybridized to 20 μg of tRNA or total *A. thaliana* RNA at 45°C and digested with 1000 units/mL of ribonuclease T_1 (Sigma, type IV, 550,500 units/mg of protein) and 40 μg/mL of ribonuclease A (Sigma, type X-A, 110 Kunitz units/mg of protein) for 1 h at 30°C.

Plant Transformation

pAR137, pAR138, and pAR139 contain the *PAT1* genomic fragments indicated in Figure 4 in the binary vector pEND4K (20). Each of these plasmids was introduced into *Agrobacterium tumefaciens* strain LBA4404 by electroporation. The resulting strains were used to transform *trp1-1* and *trp1-100* (see below) *A. thaliana* mutants to kanamycin resistance as described (21, 33), with the following modifications. Media contained 50 μM tryptophan, and selection against the *A. tumefaciens* was with 50 mg of carbenicillin/L. Shoots were induced to form roots on rooting medium (9). Blue light-induced photochemical degradation was reduced by filtration with yellow plexiglass filters (Polycast 2208, Curbell) as described (31) until the rooted shoots were transferred to soil. The kanamycin resistance of progeny seedlings was tested on PNS medium (16) containing 25 mg of kanamycin sulfate/L.

RESULTS

Isolation and Characterization of *trp1-100*

The blue fluorescence phenotype of *trp1* plants was used to screen for additional alleles of *trp1* from an M_2 population of mutagenized wild-type *A. thaliana* Columbia ecotype seeds (M_2 indicates the progeny that result from self pollination of plants derived from mutagenized seeds [22]). One isolate, named *trp1-100,* is blue fluorescent and resistant to 5-meth-

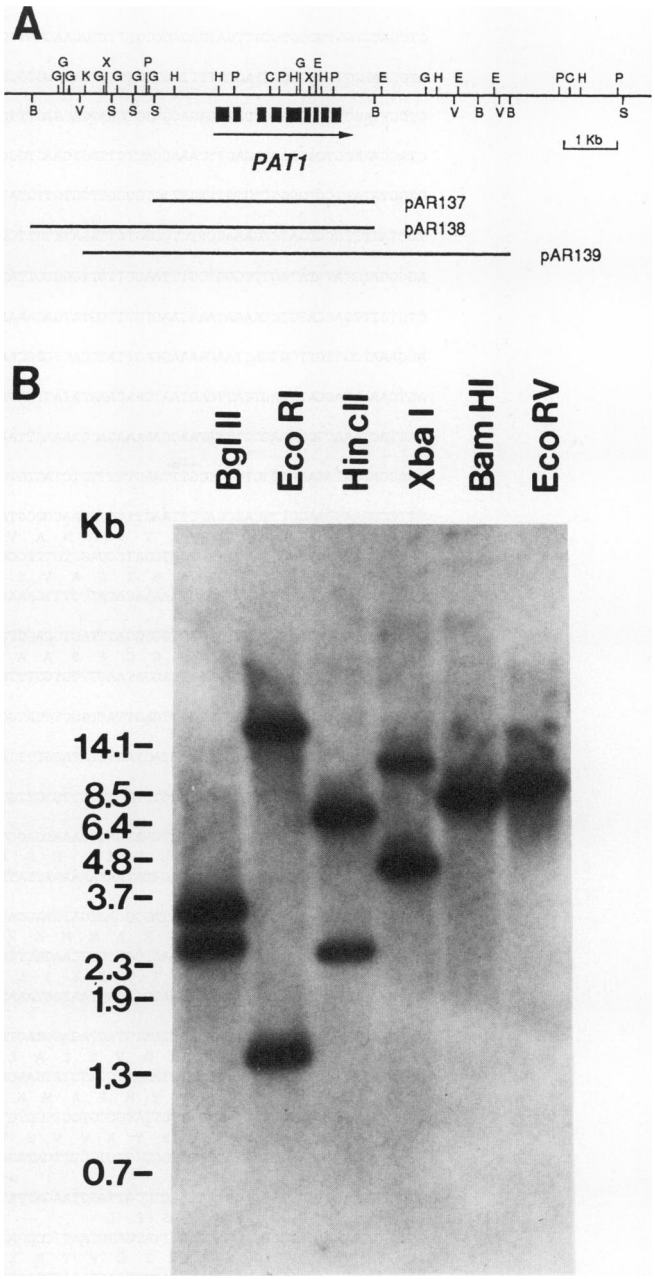

Figure 4. *PAT1* appears to be a single-copy gene. A, Restriction map of the *PAT1* gene. A *PAT1* genomic clone was mapped with the following restriction enzymes: *Bam*HI (B), *Bgl*II (G), *Eco*RI (E), *Eco*RV (V), *Hin*cII (C) (*Hin*cII also cuts at *Sal*I sites), *Hin*dIII (H), *Kpn*I (K), *Pst*I (P), *Sal*I (S), *Xba*I (X), and *Xho*I (no sites). Black boxes represent the positions of the *PAT1* exons, and the arrow indicates the direction of transcription and limits of the mature mRNA. Lines below the map show the extent of the fragments used to complement the *trp1* mutations in transgenic plants. B, Genomic Southern blot. A Southern blot of *A. thaliana* genomic DNA digested with the indicated enzymes was probed with a *PAT1* cDNA.

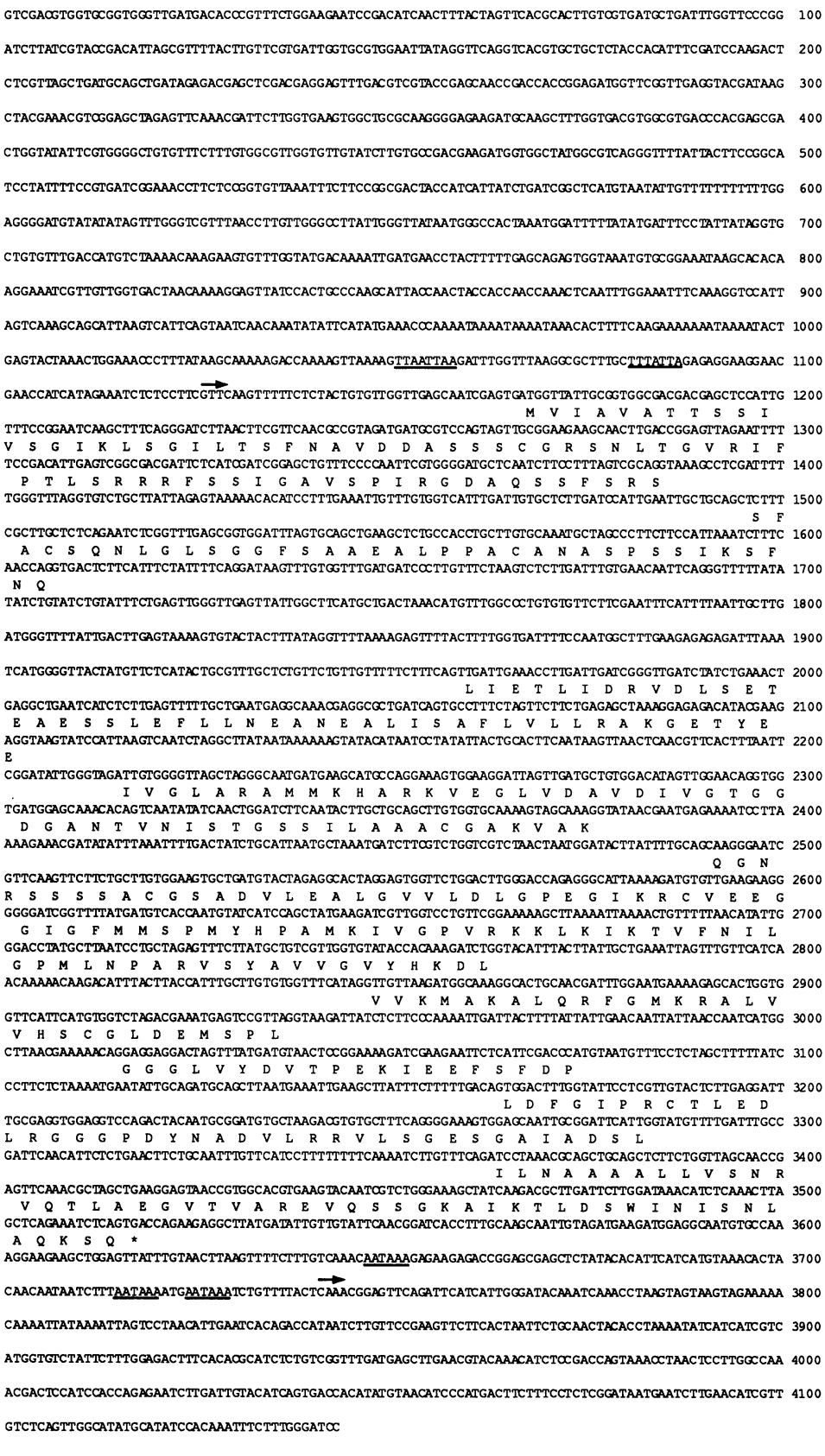

Figure 5. DNA sequence of the *PAT1* gene. The arrows indicate the major mRNA 5′ end, as determined by primer extension, and the site of poly(A) addition in pNN423. Possible TATA boxes and polyadenylation signals are underlined.

ylanthranilate, a compound that wild-type plants convert to toxic 5-methyltryptophan (22), but differs from the *trp1-1* plants in important ways. Unlike *trp1-1*, the *trp1-100* plants do not require exogenous tryptophan for growth on defined medium, and they have normal growth morphology and fertility, as shown in Figure 2. The *trp1-100* plants are less fluorescent than *trp1-1* plants, and this difference becomes more pronounced with age.

Genetic and biochemical analysis indicates that the *trp1-1* and *trp1-100* mutations are allelic. The F_1 progeny of crosses between *trp1-1* and *trp1-100* are blue fluorescent, indicating that the mutations cannot complement each other. Furthermore, all 568 of the F_2 progeny derived by self pollination of the F_1 plants are blue fluorescent, confirming that the two mutations are tightly linked. Enzyme assays showed that both *trp1-1* and *trp1-100* plants are defective in PAT activity (22, and J. Li and R. Last, unpublished results).

Isolation of the *PAT1* Gene

We sought to isolate a gene encoding PAT from *A. thaliana* because *trp1* mutant plants are known to be defective in this enzyme. *A. thaliana* cDNA clones capable of suppressing an *E. coli trpD*$^-$ PAT mutation were previously identified (13). We determined the complete double-strand sequence of the insert in pNN423 and partial sequences of several other suppressing cDNAs. These cDNAs differ in length at the 5′ end but are otherwise identical throughout the regions analyzed. We concluded that all were derived from the same gene and compiled a 1.6-kb composite cDNA sequence for further analysis.

The amino acid sequence of the only large open reading frame in this cDNA sequence (444 residues, predicted M_r of 46,517) is that of a PAT protein. Alignment of the *A. thaliana* amino acid sequence with a variety of microbial PATs is shown in Figure 3. Compared individually with each PAT enzyme, the *A. thaliana* protein has 33 to 40% identical residues and an additional 38 to 42% analogous amino acids, for a total similarity of 73 to 81%. This similarity extends over the length of the microbial proteins and includes several blocks of highly conserved amino acids. There is a single tryptophan codon in the *PAT1* coding region, consistent with the scarcity of tryptophan residues in the tryptophan biosynthetic enzymes of most organisms (8). The *A. thaliana* protein has an amino-terminal extension of approximately 105 amino acids that is not found in the microbial proteins. This region has properties common to plastid leader peptides (34), including an abundance of serine and threonine residues and a net positive charge. Based on the ability to suppress the *E. coli* mutation and the similarity in predicted protein sequence, we named the gene from which the cDNAs were derived *PAT1*.

Evidence for a Single *PAT* Gene

The *PAT1* cDNA was used as a probe for *A. thaliana* genomic Southern blot analysis. As shown in Figure 4B, the only fragments that hybridize are those expected for the *PAT1* gene. A single band is seen when *A. thaliana* DNA is digested with an enzyme that does not cut within *PAT1*

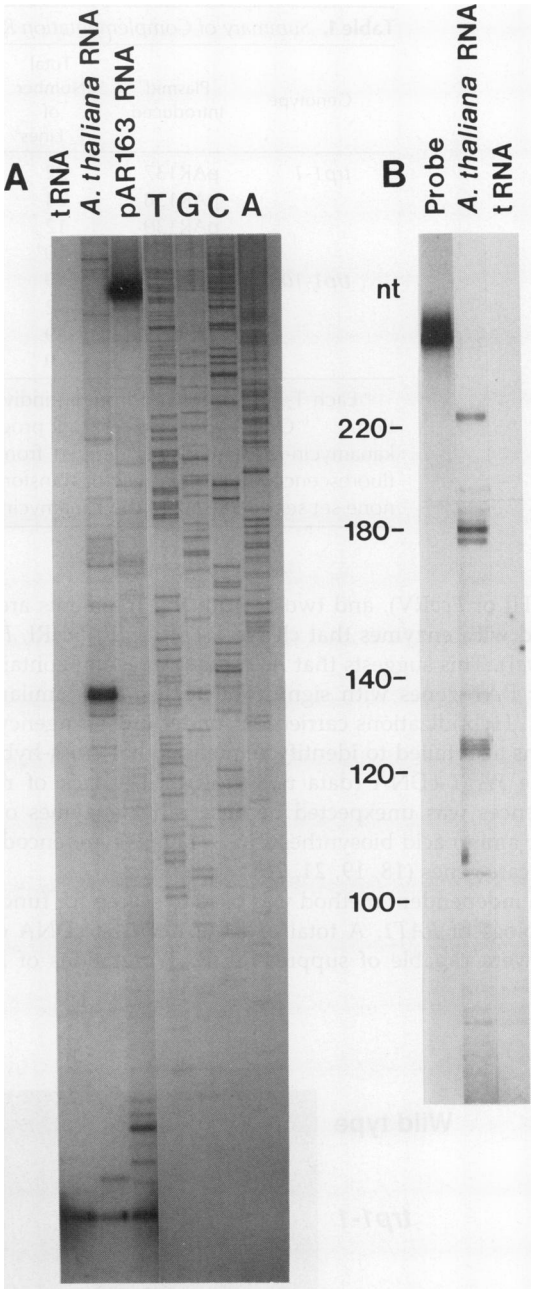

Figure 6. Mapping the 5′ end of *PAT1* mRNA. A, Primer extension analysis. Primer extension products from 5′ end-labeled ARP1 oligonucleotide annealed to 10 μg of tRNA, 10 μg of *A. thaliana* total RNA, or in vitro-synthesized *PAT1* RNA from pAR163 were separated in a denaturing 8% polyacrylamide gel. Sequencing reactions with pAR134Δ74 template and the ARP1 primer are labeled to reflect the coding strand shown in Figure 5. B, RNase protection analysis. A uniformly labeled riboprobe complementary to nucleotides 1005 to 1247 was annealed to 20 μg of *A. thaliana* total RNA or 20 μg of tRNA and digested with ribonucleases A and T_1. Protected fragments were resolved in a 6% sequencing gel and visualized by autoradiography. Undigested probe is shown from a shorter exposure of the same gel. The RNA sizes shown were estimated from a DNA sequencing ladder by correcting for the difference in migration of RNA and DNA (28).

Table I. *Summary of Complementation Results*

Genotype	Plasmid Introduced	Total Number of T_2 Lines[a]	Number Analyzed	Kanamycin-Resistant Lines		
				Total	Nonfluorescent[b]	Fluorescent[b]
trp1-1	pAR137	22	8	8	7	1[c]
	pAR138	24	8	8	8	0
	pAR139	12	7	6	6	0
	pEND4K	0[d]				
trp1-100	pAR137	9	9	8	7	1[c]
	pAR138	2	2	1	1	0
	pAR139	9	6	5	4	1[e]
	pEND4K	9	7	6	0	6[e]

[a] Each T_2 line is derived from an individual regenerated primary transformant (T_1) by self pollination. [b] Only kanamycin-resistant progeny were scored for blue fluorescence. All nonfluorescent, kanamycin-resistant plants derived from *trp1-1* were also prototrophic. [c] Segregated for blue fluorescence. [d] All 62 vector-transformed putative transgenic plants were small and bushy, and none set seed. [e] All of the kanamycin-resistant plants were blue fluorescent.

(*Bam*HI or *Eco*RV), and two hybridizing fragments are generated with enzymes that cleave *PAT1* (*Bgl*II, *Eco*RI, *Hin*cII, or *Xba*I). This suggests that *A. thaliana* does not contain any other PAT genes with significant nucleic acid similarity to *PAT1*. Hybridizations carried out under low stringency conditions also failed to identify sequences that cross-hybridize to the *PAT1* cDNA (data not shown). The lack of related sequences was unexpected because other enzymes of aromatic amino acid biosynthesis in *A. thaliana* are encoded by duplicate genes (18, 19, 21, 26).

An independent method was used to search for functional homologs of *PAT1*. A total of 81 *A. thaliana* cDNA clones that were capable of suppressing PAT mutations of *E. coli* were isolated as described (13) and screened by colony hybridization or restriction enzyme analysis. All 81 cDNAs hybridize strongly to the *PAT1* cDNA or contain a common internal 0.8-kb *Pst*I restriction fragment that is present in the *PAT1* cDNAs. These results suggest that the cDNAs were all derived from the *PAT1* gene. Our working hypothesis is that *PAT1* is a single-copy gene, based on the lack of evidence for any other gene encoding PAT.

Characterization of the *PAT1* Gene

The *PAT1* cDNA was used as a probe to isolate genomic clones from an *A. thaliana* lambda library. The restriction

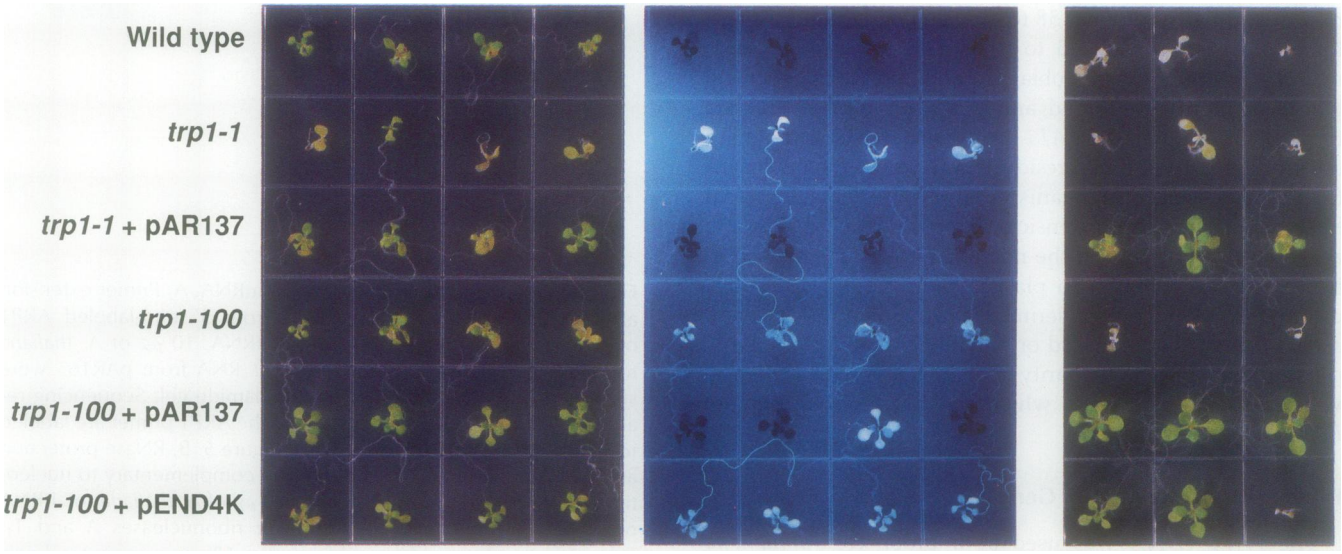

Figure 7. *PAT1* complementation of *trp1*. Wild-type (top row), *trp1–1* (second row), *trp1–1* transformed with pAR137 (third row), *trp1–100* (fourth row), *trp1–100* transformed with pAR137 (fifth row), and *trp1–100* transformed with pEND4K (bottom row) plants were grown on PNS or PNS containing 50 μM tryptophan and 25 mg/L of kanamycin, as indicated, for 10 d. The plants were photographed under visible or shortwave UV light (center panel). Some blue-fluorescent, kanamycin-sensitive individuals can be seen in segregating transgenic lines.

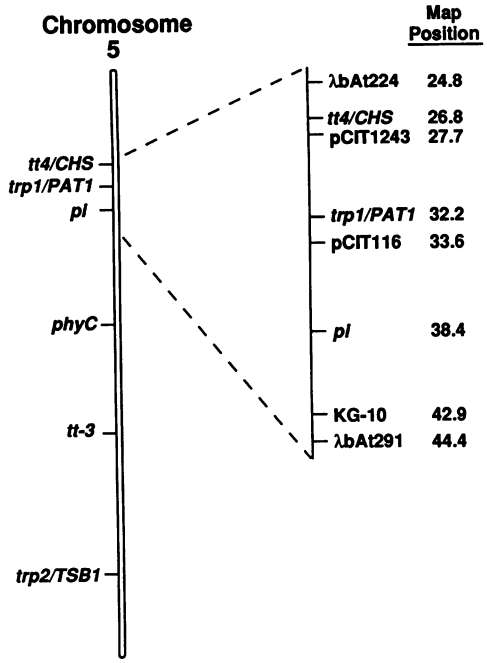

Figure 8. Map position of *PAT1*. The map location of the *PAT1* gene was determined by RFLP mapping (in the laboratory of E. Meyerowitz). Genetic crosses confirmed that *trp1/PAT1* and *tt4/CHS* are closely linked.

map of the *PAT1* region in the recombinant phage λAR1–5 is shown in Figure 4A. The DNA sequence of the 4.1-kb *Sal*I-*Bam*HI fragment containing the *PAT1* gene is shown in Figure 5. Comparison with the cDNA sequence indicates that the *PAT1* gene consists of nine exons 60 to 273 bp long, separated by eight introns ranging in size from 78 to 352 bp. Each intron is bounded by conserved GT and AG dinucleotides (2).

The 5′ ends of the *PAT1* mRNA were mapped by primer extension and RNase protection experiments. The major primer extension product, indicated by the arrowhead in Figure 6A, maps to position 1126, 41 nucleotides upstream of the *PAT1* ATG. To test the possibility that this product is an artifact caused by premature dissociation of the reverse transcriptase, an in vitro-synthesized control RNA was used as a template. pAR163 RNA is the same sense strand as PAT1 mRNA but contains an additional 121 nucleotides of PAT1 sequence upstream from position 1126. Full extension from the ARP1 primer annealed to pAR163 RNA should yield a 265-nucleotide product. As shown in Figure 6A, reverse transcriptase synthesizes the expected full-length fragment from pAR163 RNA template, suggesting that the major band in the *A. thaliana* RNA template reaction is derived from the 5′ end of the *PAT1* mRNA.

These 5′ end mapping results were confirmed by RNase protection experiments. A uniformly labeled RNA probe complementary to nucleotides 1005 to 1247 was hybridized to *A. thaliana* total RNA and digested with ribonucleases A and T_1. Four clusters of fragments were protected from digestion (Fig. 6B). The products 122 to 125 nucleotides long are in agreement with a *PAT1* mRNA 5′ end at position 1126. The additional fragments of approximately 105, 175 to 180, and 225 nucleotides indicate potential 5′ ends at positions 1023, 1068 to 1073, and 1143. No major primer extension products map to these positions. It is unlikely that any *PAT1* mRNA initiates upstream of the region homologous to the probe because RNA blot hybridization analysis (not shown) reveals that mature *PAT1* mRNA is approximately 1.8 kb in length. The ATG at position 1167 is almost certainly the initiation codon of the *PAT1* gene because it is the only methionine codon between 1005 and 1250.

PAT1 Is the Gene Defective in *trp1* Mutant Plants

Confirmation that *PAT1* is the gene defective in *trp1* mutant plants was sought by asking whether *PAT1* can complement *trp1* mutations. Three different *PAT1* genomic fragments (shown in Fig. 4A) were introduced into both *trp1-1* and *trp1-100* plants by *A. tumefaciens*-mediated transformation, with selection for kanamycin resistance conferred by the *npt*II gene present in the vector. Kanamycin-resistant progeny, derived by self pollination of the original transformed plants, were screened for blue fluorescence and tryptophan auxotrophy. The presence of the *PAT1* transgene was confirmed by genomic Southern blot analysis (data not shown).

Table II. *Genetic Linkage of trp1-100 and tt4*

Phenotypes of blue fluorescent (*trp1-100/trp1-100*) F_2 plants are shown.

Recessive Marker	Phenotype[a]		χ^2[b]
	Mutant	Wild-type	
Cross 1. *tt4/tt4; er/er* × *trp1-100/trp1-100; gl1/gl1*			
tt4	0	121	n.a.
er	32	89	0.135
gl1	31	90	0.025
Cross 2. *tt5/tt5; er/er* × *trp1-100/trp1-100; gl1/gl1*			
tt5	8	32	0.533
er	12	28	0.533
gl1	12	28	0.533

[a] *tt4/tt4* and *tt5/tt5* plants have yellow seeds. *er/er* plants have compact rosettes, short petioles, and blunt fruits. *gl1/gl1* plants lack trichomes. [b] Based on an expected mutant phenotype frequency of 25%. $P > 0.05$ for all. n.a., Not applicable.

The *PAT1* transgenes complement both *trp1* mutations. As summarized in Table I and Figure 7, transgenic plants containing any of the three *PAT1* fragments are no longer blue fluorescent. The *trp1-1* plants became prototrophic with wild-type morphology and fertility upon introduction of the *PAT1* gene. In contrast, transformation with the pEND4K vector did not reverse the *trp1* phenotypes. All 62 vector-transformed *trp1-1* plants had the bushy, infertile phenotype of the *trp1-1* parents, precluding progeny analysis. The kanamycin-resistant progeny of *trp1-100* plants transformed with pEND4K remained blue fluorescent, showing that the loss of fluorescence was not due to the transformation procedure.

The complementing activity in the transformed plants is genetically linked to kanamycin resistance. Many transformants give rise to roughly one-quarter blue-fluorescent and kanamycin-sensitive progeny, as expected if the plants are heterozygous for the dominant *PAT1* and *npt*II genes at a single locus. All of the blue-fluorescent progeny in these segregating populations are kanamycin sensitive, and none of the kanamycin-resistant plants are blue fluorescent. This demonstrates the tight linkage between the *PAT1* and *npt*II genes expected if both genes are contiguous in the transferred DNA. It also confirms that the *PAT1* gene is correcting all of the phenotypes associated with both the *trp1-1* and *trp1-100* mutations, rather than a *trp1* reversion or a tissue culture artifact.

RFLP mapping was also used to support the idea that *PAT1* is the gene defective in *trp1* mutant plants. The Columbia (Col-0) and Landsberg *erecta* (Ler) ecotypes of *A. thaliana* differ in the length of an *Eco*RV fragment that hybridizes to the *PAT1* gene (3.9 kb in Col-0, 4.8 kb in Ler). The *trp1-100*$_{Col-0}$ plants were crossed to TRP1$_{Ler}$ plants, and subsequent self-crossed generations were analyzed. After scoring the plants for blue fluorescence, genomic DNA was isolated from F_3 progeny of 30 individual F_2 plants. These DNAs were digested with *Eco*RV and subjected to Southern blot analysis using the *PAT1* gene as a probe. In all cases, the *TRP1* locus cosegregated with the *PAT1* RFLP; the Col-0 3.9-kb fragment was found in all blue-fluorescent plants, the Ler 4.8-kb fragment was in the true-breeding nonfluorescent plants, and both bands were present in heterozygous populations.

The linkage of the *trp1* mutation and the *PAT1* gene was confirmed by independent genetic means. RFLP mapping (performed in the laboratory of E. Meyerowitz) placed the *PAT1* gene at position 32.2 on chromosome 5, 5.4 map units from the *tt4* mutation (or *CHS* gene, see Fig. 8). If the *trp1* mutations are in the *PAT1* gene, we would expect very few recombinants from a cross between *tt4* and *trp1* plants.

To test this hypothesis, blue-fluorescent F_2 plants from a *tt4* × *trp1-100* cross were scored for yellow seeds, characteristic of homozygous *tt4* plants. If *trp1-100* is 5.4 map units from *tt4*, we expect one in every 340 *trp1-100* homozygotes to be homozygous for *tt4*. No *tt4/tt4* plants were found out of 121 blue-fluorescent (*trp1-100/trp1-100*) F_2 plants screened (Table II), confirming that *trp1-100* is linked to *tt4*, and supporting the idea that the *trp1-100* mutation is in the *PAT1* gene. The other recessive markers in this cross, *gl1* and *er*, segregated as expected for unlinked genes (Table II). The absence of *trp1-100; tt4* double homozygotes could be due to a synthetic lethal interaction between flavonoid (*tt*) and tryptophan biosynthetic mutations rather than genetic linkage. This hypothesis was tested in a control cross between *tt5*, an unlinked flavonoid biosynthetic mutation on chromosome 3 (29), and *trp1-100* plants. Approximately one-quarter of the blue-fluorescent F_2 plants had yellow seeds (Table II), as expected for unlinked markers.

DISCUSSION

Here we report the isolation of the *A. thaliana PAT1* gene and show by complementation and mapping that it is the gene defective in *trp1* mutant plants.

Amino acid sequence comparisons confirm that *PAT1* encodes a PAT. The most highly conserved block of amino acids (residues 181–189 of the *A. thaliana* sequence) has been implicated by mutational analysis as a phosphoribosyl pyrophosphate binding site (10), and is similar to the Walker consensus sequence for nucleotide binding (35). Although most of the highly conserved amino acids in microbial PATs are also found in the *A. thaliana* protein, a few are not. The glutamine, methionine, and glycine residues at positions 213, 285, and 336 of the *A. thaliana* sequence are histidine, leucine, and threonine residues, respectively, in the other PATs shown in Figure 2.

The presence of a putative plastid leader sequence suggests that the mature PAT protein is localized to the chloroplast, the presumed site of tryptophan synthesis (3, 30). Other enzymes in the aromatic amino acid pathways, such as 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase, 5-*enol*pyruvylshikimate-3-phosphate synthase, and tryptophan synthase β, have similar leader sequences and were shown biochemically to be present in chloroplasts (11, 14, 21, 25).

Not all of the cDNAs analyzed encode the full leader sequence. Although we cannot rule out that these clones were derived from functional mRNAs that encode nonplastid targeted enzymes, we favor the hypothesis that these are incomplete cDNAs. The first in-frame initiation codon in these cDNAs is at nucleotide 2236 in exon 4. Thus, these clones would encode a truncated PAT enzyme that is missing a significant region of similarity to the microbial enzymes. The translational initiation of these cDNA clones in *E. coli* is almost certainly mediated by an ATG adjacent to the insert in the λYES vector (13).

In many microorganisms, the enzymes catalyzing two different reactions in the tryptophan pathway are fused into bifunctional proteins (7). For example, the *E. coli trpD* gene encodes a fusion of PAT and the small subunit of anthranilate synthase. There is no evidence that *PAT1* or any of the other *A. thaliana* aromatic amino acid genes analyzed to date encode fused proteins. The observation that the *A. thaliana PAT1* cDNA can suppress an *E. coli trpD*$^-$ mutation that causes a defect only in transferase activity illustrates that the transferase and synthase activities do not need to be fused into a single protein to function in *E. coli*. Full-length *PAT1* cDNAs are effective at suppressing the *trpD*$^-$ mutation, indicating that the putative leader peptide either does not interfere with protein function or is removed by *E. coli*. The ability of plant proteins to function in bacteria or yeast greatly facilitates the isolation of plant genes for which analogous

microbial mutations exist and will facilitate the analysis of functional domains of the plant proteins.

PAT1 seems to be the exception to the rule that *A. thaliana* aromatic amino acid biosynthetic enzymes are encoded by more than one gene. Another enzyme that is apparently encoded by a single gene in *A. thaliana* is acetolactate synthase (23), which catalyzes the first common step in the biosynthesis of leucine, isoleucine, and valine. There are several unlikely ways in which our screen for a second *PAT* gene could have failed. A widely divergent but functionally redundant gene that is poorly represented in the cDNA library or is nonfunctional in *E. coli* might have been missed in a screen for cDNAs encoding PAT activity. Alternatively, a recent duplication may have resulted in two *PAT*-containing chromosomal segments over 14 kb in length with identical restriction maps for the seven enzymes tested. The ability to isolate recessive *trp1* mutations suggests that *PAT1* is single copy, or that a second gene, if present, either provides very little PAT activity or is expressed in different tissues than *PAT1*.

The *trp1-1* and *trp1-100* plants differ in growth habit, fertility, and nutritional requirements. The fact that *PAT1* complements all tested phenotypes of both mutants confirms the genetic evidence that these very different mutants have lesions in the same gene. The *trp1-100* plants are much easier to work with than *trp1-1* plants because even though they are blue fluorescent they grow normally and set seed well. The *trp1-1* and *trp1-100* mutations are probably missense mutations that differ in the degree to which they impair PAT function. Experiments to test this hypothesis are in progress.

The *PAT1* gene, in conjunction with the *trp1* mutants, should prove to be a useful tool for the study of *A. thaliana*. Many of the techniques that allow sophisticated modifications of yeast, microbial, and even mammalian genomes are unavailable in plant molecular genetics. The antibiotic resistance genes currently employed have some disadvantages, such as allowing a relatively high number of individuals lacking the genes to escape selection. The small number of useful selectable markers currently available restricts the number of manipulations possible in a single strain. The *PAT1* gene could serve as a marker in plant transformation, allowing selection in *trp1-1* auxotrophs or screening for the loss of blue fluorescence in *trp1-100* plants. *PAT1* might also prove useful as a gene that can be selected against, as both *trp1-1* and *trp1-100* plants are resistant to 5-methylanthranilate (22, data not shown). Selection against a gene has been used to increase the recovery of transformants in which a DNA fragment has integrated by homologous recombination (4).

Genes whose expression can be monitored visually, such as *lacZ* in *E. coli* or *ADE2* in *Saccharomyces cerevisiae*, have proven valuable in a variety of applications such as genetic screens and recombination studies. The bacterial *uidA* gene, encoding β-glucuronidase, is commonly used as a visible reporter to study tissue-specific promoter activity in plants, but detection usually requires a treatment that kills the sample. The *PAT1* gene might serve as a visible reporter of gene expression because the presence or activity of *PAT1* can be easily detected in whole living plants by screening for fluorescence.

ACKNOWLEDGMENTS

We thank John Mulligan (Stanford University) for generously providing the λYES *A. thaliana* cDNA library and plasmid pNN423. We are also grateful to Leonard Medrano and Elliott Meyerowitz (California Institute of Technology) for RFLP mapping pAR119. We would like to thank Lesilee Rose, Brian Keith, Scott Leisner, and Deborah Mahoney for helpful comments on the manuscript. The technical assistance of Andrea Barczak and photographic help of Tana Ebaugh of the Boyce Thompson Institute are gratefully acknowledged.

LITERATURE CITED

1. **Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K** (1992) Current Protocols in Molecular Biology. Wiley Interscience, New York
2. **Breathnach R, Chambon P** (1981) Organization and expression of eucaryotic split genes coding for proteins. Annu Rev Biochem **50:** 349–383
3. **Bryan JK** (1990) Advances in the biochemistry of amino acid biosynthesis. *In* BJ Miflin, PJ Lea, eds, The Biochemistry of Plants, Vol 16. Academic Press, New York, pp 161–195
4. **Capecchi MR** (1989) The new mouse genetics: altering the genome by gene targeting. Trends Genet **5:** 70–76
5. **Church GM, Gilbert W** (1984) Genomic sequencing. Proc Natl Acad Sci USA **81:** 1991–1995
6. **Cone K** (1989) Yet another plant DNA prep. Maize Genet Newsletter **63:** 68
7. **Crawford IP** (1987) Synthesis of tryptophan from chorismate: comparative aspects. Methods Enzymol **142:** 293–300
8. **Crawford IP** (1989) Evolution of a biosynthetic pathway: the tryptophan paradigm. Annu Rev Microbiol **43:** 567–600
9. **Damm B, Willmitzer L** (1988) Regeneration of fertile plants from protoplasts of different *Arabidopsis thaliana* genotypes. Mol Gen Genet **213:** 15–20
10. **de Boer JG, Glickman BW** (1991) Mutational analysis of the structure and function of the adenine phosphoribosyltransferase enzyme of Chinese hamster. J Mol Biol **221:** 163–174
11. **della-Ciopa G, Bauer SC, Klein BK, Shah DM, Fraley RT, Kishore GM** (1986) Translocation of the precursor of 5-*enol*-pyruvylshikimate-3-phosphate synthase into chloroplasts of higher plants *in vitro*. Proc Natl Acad Sci USA **83:** 6873–6877
12. **DeVries S, Hoge H, Bisseling T** (1988) Isolation of total and polysomal RNA from plant tissues. *In* SB Gelvin, RA Schilperoort, DPS Verma, eds, Plant Molecular Biology Manual, B6. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 1–13
13. **Elledge SJ, Mulligan JT, Ramer SW, Spottswood M, Davis RW** (1991) λYES: a multifunctional cDNA expression vector for the isolation of genes by complementation of yeast and *Escherichia coli* mutations. Proc Natl Acad Sci USA **88:** 1731–1735
14. **Ganson RJ, D'Amato TA, Jensen RA** (1986) The two-isozyme system of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase in *Nicotiana sylvestris* and other higher plants. Plant Physiol **82:** 203–210
15. **Haughn GW, Davin L, Giblin M, Underhill EW** (1991) Biochemical genetics of plant secondary metabolites in *Arabidopsis thaliana*. The glucosinolates. Plant Physiol **97:** 217–226
16. **Haughn GW, Somerville CR** (1986) Sulfonylurea resistant mutants of *Arabidopsis thaliana*. Mol Gen Genet **204:** 430–434
17. **Henikoff S** (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351–359
18. **Keith B, Dong X, Ausubel FM, Fink GR** (1991) Differential induction of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase genes in *Arabidopsis thaliana* by wounding and pathogenic attack. Proc Natl Acad Sci USA **88:** 8821–8825
19. **Klee HJ, Muskopf YM, Gasser CS** (1987) Cloning of an *Arabidopsis thaliana* gene encoding 5-enolpyruvylshikimate-3-phosphate synthase: sequence analysis and manipulation to obtain glyphosate-tolerant plants. Mol Gen Genet **210:** 437–442

20. **Klee HJ, Yanofsky MF, Nester EW** (1985) Vectors for transformation of higher plants. Bio/Technology **3:** 637–642
21. **Last RL, Bissinger PH, Mahoney DJ, Radwanski ER, Fink GR** (1991) Tryptophan mutants in *Arabidopsis:* the consequences of duplicated tryptophan synthase β genes. Plant Cell **3:** 345–358
22. **Last RL, Fink GR** (1988) Tryptophan-requiring mutants of the plant *Arabidopsis thaliana*. Science **240:** 305–310
23. **Mazur BJ, Chui C-F, Smith JK** (1987) Isolation and characterization of plant genes coding for acetolactate synthase, the target enzyme for two classes of herbicides. Plant Physiol **85:** 1110–1117
24. **Moffatt BA, McWhinnie EA, Burkhart WE, Pasternak JJ, Rothstein SJ** (1992) A complete cDNA for adenine phosphoribosyl transferase from *Arabidopsis thaliana*. Plant Mol Biol **18:** 653–662
25. **Mousdale DM, Coggins JR** (1985) Subcellular localization of the common shikimate-pathway enzymes in *Pisum sativum* L. Planta **163:** 241–249
26. **Niyogi KK, Fink GR** (1992) Two anthranilate synthase genes in *Arabidopsis:* defense-related regulation of the tryptophan pathway. Plant Cell **4:** 721–733
27. **Pearson WR, Lipman DJ** (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA **85:** 2444–2448
28. **Sambrook J, Fritsch EF, Maniatis T** (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
29. **Shirley BW, Hanley S, Goodman HM** (1992) Effects of ionizing radiation on a plant genome: analysis of two *Arabidopsis* transparent testa mutations. Plant Cell **4:** 333–347
30. **Singh BK, Siehl DL, Connelly JA** (1991) Shikimate pathway: why does it mean so much to so many? *In* BJ Miflin, ed, Oxford Surveys of Plant Molecular and Cell Biology, Vol 7. Oxford University Press, Oxford, UK, pp 143–186
31. **Stasinopoulos TC, Hangarter RP** (1990) Preventing photochemistry in culture media by long-pass light filters alters growth of cultured tissues. Plant Physiol **93:** 1365–1369
32. **Tsuji J, Jackson EP, Gage DA, Hammerschmidt R, Somerville SC** (1992) Phytoalexin accumulation in *Arabidopsis thaliana* during the hypersensitive reaction to *Pseudomonas syringae* pv *syringae*. Plant Physiol **98:** 1304–1309
33. **Valvekens D, Van Lijsebettens M, Van Montagu M** (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. Proc Natl Acad Sci USA **85:** 5536–5540
34. **von Heijne G, Hirai T, Klosgen R-B, Steppuhn J, Bruce B, Keegstra K, Herrmann R** (1991) CHLPEP—a database of chloroplast transit peptides. Plant Mol Biol Rep **9:** 104–126
35. **Walker JE, Saraste M, Runswick MJ, Gay NJ** (1982) Distantly related sequences in the α- and β-subunits of ATP synthase, myosin kinases, and other ATP requiring enzymes and a common nucleotide binding fold. EMBO J **1:** 945–951
36. **Wright AD, Sampson MB, Neuffer MG, Michalczuk L, Slovin JP, Cohen JD** (1991) Indole-3-acetic acid biosynthesis in the mutant maize *orange pericarp*, a tryptophan auxotroph. Science **254:** 998–1000